

# On well-separated sets and fast multipole methods

Stefan Engblom\*

Department of Numerical Analysis,  
School of Computer Science and Communication,  
Royal Institute of Technology, S-100 44 Stockholm, Sweden

---

## Abstract

The notion of well-separated sets is crucial in fast multipole methods as the main idea is to approximate the interaction between such sets via cluster expansions. We revisit the one-parameter multipole acceptance criterion in a general setting and derive a relative error estimate. This analysis benefits asymmetric versions of the method, where the division of the multipole boxes is more liberal than in conventional codes. Such variants offer a particularly elegant implementation with a balanced multipole tree, a feature which might be very favorable on modern computer architectures.

**Keywords:** fast multipole method, balanced tree, asymmetric adaptive mesh, error analysis, Cartesian expansion.  
**2010 MSC:** 65M15, 65M80.

---

## 1. Motivation

We consider in this paper a general error analysis and some implementation issues for fast multipole methods (FMMs). Since their first appearance in [3, 7], these tree-based algorithms have become important computational tools for evaluating pairwise interactions of the type

$$\Phi(x_i) = \sum_{j=1, j \neq i}^N G(x_i, x_j), \quad x_i \in \mathbf{R}^D, \quad i = 1 \dots N. \quad (1.1)$$

FMMs offer an  $O(N)$ -complexity and an *a priori* error estimate, but implementing a fully fledged adaptive FMM in 3D is a daunting task [4]. Parallelization issues complicate matter even further and call for a balance between, on the one hand theoretical efficiency, and on the other hand software complexity [14, 16]. A major inconvenience with *adaptive* versions is the complicated memory access pattern which is due to the communication between levels in the multipole tree. Although this can be mitigated either through post-balancing algorithms [15], or by employing more advanced data-structures [9, Chap. 6.6], with modern data-parallel architectures it has in fact been suggested that uniform versions offer better performance [10].

An alternative is to use *asymmetric* adaptive meshes as outlined in Figures 1.1 and 4.1. Here the tree becomes balanced at the cost of a variable, but local, communication stencil. Also, the actual form of the *multipole acceptance criterion* becomes more critical as the mesh loses regularity. Comparisons between different criteria for cell-to-particle methods are found in [12], while a careful worst-case analysis of uniform FMMs is found in [11].

The main contributions of the current paper are found in Sections 2 and 3 where the precise statement of the multipole acceptance criterion and the required assumptions on the kernel  $G$  are discussed together with a general

---

\*Phone: +46-18-471 27 54, Fax: +46-18-51 19 25

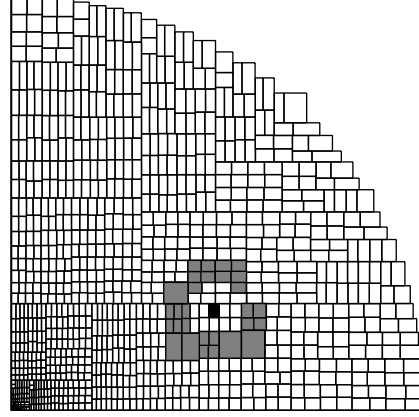
Email address: stefane@it.uu.se (Stefan Engblom)

URL: <http://user.it.uu.se/~stefane> (Stefan Engblom)

<sup>1</sup>Current address: Division of Scientific Computing, Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden.

error analysis. Our approach is inspired by the treatment in the monograph [8], but we are able to offer several important corrections. In Sections 4 and 5 implementation issues are highlighted and we also perform numerical experiments illustrating the sharpness of the theory and the efficiency of the proposed approach.

Figure 1.1: Sample asymmetric adaptive multipole mesh in 2D. The discretization is obtained by successively splitting boxes along each coordinate axis in such a way that the number of sources in the four resulting boxes is very nearly equal. In this example the grey boxes are to interact through cluster-to-cluster interactions with the black box. That is, they satisfy the  $\theta$ -criterion ( $\theta = 1/2$ ).



## 2. Well-separated sets and kernel assumptions

We begin by stating the one-parameter criterion for two sets being well-separated. This is also a correction to the too weak criterion found in [8, Eq. (8.31), p. 379].

**Criterion 2.1** ( $\theta$ -criterion). Let the sets  $S_1, S_2 \subset \mathbf{R}^D$  be contained inside two disjoint spheres such that  $\|S_1 - x_0\| \leq r_1$  and  $\|S_2 - y_0\| \leq r_2$ . Given  $\theta \in (0, 1)$ , if  $d \equiv \|x_0 - y_0\|$ ,  $R \equiv \max\{r_1, r_2\}$ , and  $r \equiv \min\{r_1, r_2\}$ , then the two sets are *well-separated* whenever  $R + \theta r \leq \theta d$ .

In other words, any of the two sets may be expanded by a factor of  $1/\theta$  and arbitrarily rotated about its center point without touching the other set. When regarded as a parameter inside the FMM it will be evident that a smaller  $\theta$  yields a smaller error at the cost of a larger communication stencil.

We remark that the  $\theta$ -criterion under consideration is symmetric in its two arguments as a reflection upon the fact that we mainly consider FMMs where the sources and the evaluation points are “roughly” the same. — Indeed, our implementation as outlined in Section 4 produces a representation of the total field in the *whole* enclosing box under consideration. An algorithm which is adaptive in sources and evaluation points separately is described in [9, Chap. 6.6; see also Fig. 5.1].

We shall need the following two simple consequences of the  $\theta$ -criterion: since  $\theta < 1$  we get

$$\frac{d + r + R}{d - R} \leq \frac{d + r + R/\theta}{d - \theta(d - r)} \leq \frac{2d}{(1 - \theta)d + \theta r} \leq \frac{2}{1 - \theta}. \quad (2.1)$$

Furthermore, writing the  $\theta$ -criterion as  $r \leq \theta(d - R) + (\theta - 1)(R - r)$ , we also see that

$$\frac{r}{d - R} \leq \frac{R}{d - r} \leq \theta. \quad (2.2)$$

For  $\alpha$  and  $\beta$  multi-indices in the normed space  $(\mathbf{Z}_+^D, |\cdot|)$  we write for brevity  $G_{\alpha\beta}(x, y) = \partial^\alpha x \partial^\beta y G(x, y)$  and additionally define factorials and powers in the usual intuitive manner. Similarly to [8, Eq. (8.15), p. 317] (but here with an explicit factor  $n!$ ) we assume that the kernel  $G$ ’s first few derivatives are equivalent to the harmonic potential:

**Assumption 2.2** (Kernel regularity). For  $\alpha, \beta \in \mathbf{Z}_+^D$  with  $|\alpha|, |\beta| \leq p + 1$ ,

$$|G_{\alpha\beta}(x, y)| \leq C \frac{n!}{\|x - y\|^{n+1}}, \quad (2.3)$$

where  $n \equiv |\alpha + \beta|$ . Additionally,  $G$  is positive and satisfies

$$\|x - y\|^{-1} \leq cG(x, y). \quad (2.4)$$

The analysis below takes place in  $D$ -dimensional Cartesian space. In order to avoid the worst-case bound  $\|\cdot\|_1 \leq \sqrt{D}\|\cdot\|$ , assuming some kind of symmetry of the kernel seems inevitable (see also [14]):

**Assumption 2.3** (Rotational invariance). For any rotation  $T$  of the coordinate system,

$$|G(x, y)| = |G(Tx, Ty)|. \quad (2.5)$$

In other words, we may freely rotate the coordinate system before expanding the kernel (provided of course that any expansion points are rotated as well). By choosing a  $T$  such that  $\|Tx\|_1 = \|x\|$ , results obtained below in  $\|\cdot\|_1$  will be transferred to the Euclidean norm without introducing any constants.

### 3. Analysis

Consider now two points  $x \in S_1$  and  $y \in S_2$ , where  $S_1$  and  $S_2$  satisfy the  $\theta$ -criterion. Approximating a unit source at  $y$  using a far-to-far translation ( $G \rightarrow \widehat{G}_p$ , centered at  $y_0$ ), followed by a far-to-local expansion ( $\widehat{G}_p \rightarrow \widetilde{G}_p$ , centered at  $x_0$ ), can be written as

$$\begin{aligned} G(x, y) &= \widetilde{G}_p(x, y; x_0, y_0) - [\widetilde{G}_p(x, y; x_0, y_0) - \widehat{G}_p(x, y; y_0)] - [\widehat{G}_p(x, y; y_0) - G(x, y)] \\ &=: \widetilde{G}_p(x, y; x_0, y_0) - E_{\text{far-to-local}} - E_{\text{far-to-far}}, \end{aligned} \quad (3.1)$$

with  $p$  the order of the expansion. We consider the two errors in turn.

The integral form for the remainder of the  $D$ -dimensional Taylor series becomes

$$\begin{aligned} |E_{\text{far-to-far}}| &= (p+1) \left| \sum_{|\beta|=p+1} \frac{(y-y_0)^\beta}{\beta!} \int_0^1 (1-t)^p G_{0,\beta}(x, y_0 + t(y-y_0)) dt \right| \\ &\leq C(p+1) \frac{\|y-y_0\|_1^{p+1}}{\|x-y_0\|^{p+2}} \int_0^1 \frac{(1-t)^p}{\left(1-t \frac{\|y-y_0\|}{\|x-y_0\|}\right)^{p+2}} dt, \end{aligned}$$

using (2.3) and the multinomial theorem. As discussed in conjunction with Assumption 2.3 above we may now replace  $\|y-y_0\|_1$  with  $\|y-y_0\|$ . Using the triangle inequality and (2.2) yields  $\|y-y_0\|/\|x-y_0\| \leq \theta$  so that from (2.4),

$$\frac{1}{\|x-y_0\|} \leq cG(x, y) \frac{\|x-y\|}{\|x-y_0\|} \leq cG(x, y) \left(1 + \frac{\|y-y_0\|}{\|x-y_0\|}\right) \leq cG(x, y)(1+\theta). \quad (3.2)$$

Using  $1-t \leq 1-\theta t \leq 1-t\|y-y_0\|/\|x-y_0\|$  we readily bound the integrand and get

$$|E_{\text{far-to-far}}| \leq cC(p+1) \theta^{p+1} \frac{1+\theta}{1-\theta} G(x, y). \quad (3.3)$$

As for the first term in (3.1) we obtain this time a *sum* of integral remainders,

$$\begin{aligned} |E_{\text{far-to-local}}| &= (p+1) \left| \sum_{\substack{|\alpha|=p+1 \\ |\beta| \leq p}} \frac{(x-x_0)^\alpha}{\alpha!} \frac{(y-y_0)^\beta}{\beta!} \int_0^1 (1-t)^p G_{\alpha,\beta}(x_0 + t(x-x_0), y_0) dt \right| \\ &\leq C(p+1) \sum_{k=0}^p \frac{\|x-x_0\|^{p+1}}{(p+1)!} \frac{\|y-y_0\|^k}{k!} \int_0^1 \frac{(1-t)^p (k+p+1)!}{\|x_0 + t(x-x_0) - y_0\|^{k+p+2}} dt, \end{aligned}$$

where the multinomial theorem and the rotational invariance were used twice. The sum can be evaluated when the upper limit tends to  $\infty$ ; hence by uniform convergence,

$$\begin{aligned} |E_{\text{far-to-local}}| &\leq C(p+1)\|x-x_0\|^{p+1} \int_0^1 \frac{(1-t)^p}{(\|x_0+t(x-x_0)-y_0\|-\|y-y_0\|)^{p+2}} dt \\ &\leq C(p+1) \frac{\|x-x_0\|^{p+1}}{(\|x_0-y_0\|-\|y-y_0\|)^{p+2}} \int_0^1 \frac{(1-t)^p}{\left(1-t \frac{\|x-x_0\|}{\|x_0-y_0\|-\|y-y_0\|}\right)^{p+2}} dt. \end{aligned}$$

Using (2.4), the triangle inequality, and (2.1) we get (compare with (3.2))

$$\frac{1}{\|x_0-y_0\|-\|y-y_0\|} \leq cG(x,y) \frac{\|x-y\|}{\|x_0-y_0\|-\|y-y_0\|} \leq cG(x,y) \frac{2}{1-\theta}. \quad (3.4)$$

For the integrand we use the same type of bound as before and finally get

$$|E_{\text{far-to-local}}| \leq cC(p+1) \theta^{p+1} \frac{2}{(1-\theta)^2} G(x,y). \quad (3.5)$$

By summing the contributions (3.3) and (3.5) from  $N$  positive potentials (cf. (2.4)) we thus conclude that the relative error for the  $p$ th order fast multipole method under the  $\theta$ -criterion is bounded by a constant  $\times \theta^{p+1}/(1-\theta)^2$ .

In the above derivation we are lead to the  $\theta$ -criterion by the initial requirement that  $\|y-y_0\|/\|x-y_0\| \leq \theta$ . Clearly, the only natural way to obtain this is from the triangle inequality and the slightly stronger requirement that  $\|y-y_0\|/(\|x_0-y_0\|-\|x-x_0\|) \leq \theta$  (cf. (2.2)). Despite this clear reasoning we have not been able to find our version of the  $\theta$ -criterion made explicit in the literature. Presumably, this is due to the fact that quadratic meshes are so popular.

We note also that the integral remainder term was consistently used instead of the Lagrangian version (hinted at also in [12, 17]). To see why, note that in bounding e.g.  $E_{\text{far-to-far}}$  above we would otherwise obtain terms of the form  $G_{0,\beta}(x, y_0 + \xi(y-y_0))$  with  $\xi \in [0, 1]$  and  $|\beta| = p+1$ . The only sensible bound now includes the factor  $(1-\theta)^{-(p+2)}$  suggesting that the convergence would deteriorate as  $\theta \rightarrow 1/2$ . By contrast, the bounds (3.3) and (3.5) are perfectly regular for any  $0 < \theta < 1$ .

It is to be stressed that we like to view the above analysis as a kind of *template*; the precise form of the assumptions and their implications are all very clearly visible. The effect of changing them can therefore readily be assessed.

#### 4. Implementation

The current paper on the  $\theta$ -criterion stems mainly from the fact that allowing an asymmetric adaptive splitting when creating the FMM mesh makes a convenient implementation possible. The ease of implementation is mainly thanks to the fact that the associated multipole tree is always balanced so that a static memory layout becomes natural. As a result, neighboring boxes are always arranged in the same level in the multipole tree, thereby facilitating the communication between them (see Figure 4.1 for an illustration).

We now give a brief description of our current implementation. For more detailed information the interested reader is kindly referred to the freely available code itself (see Section 5.1).

The multipole mesh is constructed first and is obtained by recursively splitting the source points at or near their median in each coordinate direction. A very fast and well-known algorithm is available for doing just this; this is the *median-of-three* selection algorithm which is most often used when implementing *quicksort* [13, Chap. 9]. The source points themselves are naturally stored in a *pyramid* data-structure, a  $2^d$ -*tree* cut at a certain maximum level [9, Chap. 5.3].

After the sources have been assigned a box at the finest level, the *connectivity* information is determined. At any level in the tree the boxes are either decoupled or, respectively, strongly/weakly coupled. For each box  $b$ , the strong connections  $S(p)$  of its parent box  $p$  are examined; all children of a box in  $S(p)$  that satisfy the  $\theta$ -criterion with respect to  $b$  become weakly coupled to  $b$  — the rest remain strongly coupled. This simple rule together with the fact that a box is always strongly connected to itself recursively defines the connectivity for the whole tree. The resulting topology is conveniently stored in sparse matrices with sparsity patterns known *before* each level is to be examined.

In the downward/upward phases the expansions are shifted as usual according to parent/child relations. The critical far-to-local shift follows the weak connectivity pattern and, at the finest level, the remaining strong connections are evaluated directly. As suggested in [5, 6] all shifts in the implementation tested below relies on BLAS Level 3 routines with constant transition matrices (using pre- and post-scaling according to the local geometry).

A few structural optimizations are optionally possible. For instance, a box weakly coupled to all children of a parent box can often interact via the latter instead (whenever the  $\theta$ -criterion with respect to that parent is true). This is visible in Figure 1.1 where some of the weakly coupled grey boxes are noticeable large — here the interactions are in fact managed via parent boxes. A related idea at the *finest* level is to investigate strongly coupled boxes of very different radii,  $r \ll R$ , say. If the  $\theta$ -criterion is true *when the roles of  $r$  and  $R$  are exchanged*, then the sources in the larger box can be directly converted into a near field expansion in the smaller box, and the outgoing expansion from the smaller can simply be evaluated at each point in the larger box. This optimization was suggested already in [3].

Since it is beneficial to admit as large set of boxes as possible into the set of weak connections, it is natural to try to somehow locally relax the  $\theta$ -criterion. Rather than uniformly enforcing a single value of  $\theta$  for example, *different*  $\theta_i$ 's may well be accepted provided that an estimate of the total error remains bounded. For instance, if  $n$  observed values of  $\theta_i$  satisfy  $\sum_i \theta_i^{p+1} / n \leq \theta^{p+1}$ , then the relative error estimate is still  $O(\theta^{p+1})$ . Ideas along these lines are discussed in [12, 18].

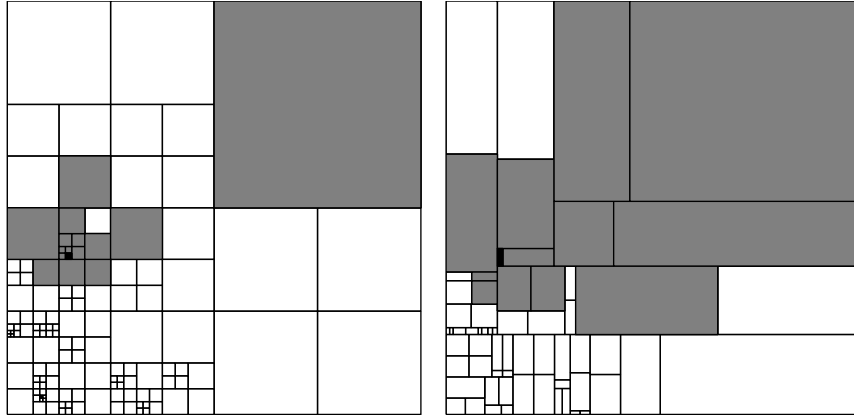


Figure 4.1: Standard midpoint adaptivity (*left*) vs. asymmetric adaptivity (*right*) for the same set of point sources (not shown). In both cases the boxes that are strongly connected to the black box are shaded grey ( $\theta = 1/2$ ). In the conventional approach, strongly connected boxes of different sizes imply cross-level communication in the multipole tree. By splitting the boxes at the median point source rather than at their geometrical midpoints, those boxes always belong to the same level.

We conclude this section by very briefly discussing the algorithmic complexity. It is known that there typically exists extremely non-uniform distributions of  $N$  source points that make most tree-codes (even the adaptive ones) run in time proportional to  $O(N^2)$  [1]. Practical experience is usually much better, although for certain applications, simpler algorithms are occasionally preferred (see [2], and further the discussions in [8, Chap. 8.7], and [9, Chap. 6.6.3]).

In any case, the serial complexity for a 2D implementation using asymmetric adaptivity can be estimated to be roughly proportional to  $\theta^{-2} p^2 \cdot N$ , since each of about  $N$  boxes at the finest level is to interact through cluster-to-cluster interactions with on the order of  $\theta^{-2}$  boxes at the same level, and since each such shift requires on the order of  $p^2$  operations. For a given relative tolerance TOL, the analysis in Section 3 implies  $p \sim \log \text{TOL} / \log \theta$ , so that the complexity is  $O(\theta^{-2} \log^{-2} \theta \cdot N \log^2 \text{TOL})$ , thus indicating that the choice  $\theta = \exp(-1) \approx 0.368$  is nearly optimal. In practice, the best value of  $\theta$  as well as the optimal number of subdivisions is rather dependent on the hardware and should be determined from experience. On balance we have found that the convenient choice  $\theta = 0.5$  and subdividing the points until the number of sources per box is  $\sim 20$  works very well in practice.

## 5. Experiments

As illustrations to the analysis in Section 3 and in order to highlight some of the benefits with the proposed adaptivity we report some results from our two-dimensional implementation which employs the classical complex polynomial/multipole representation [3, 7].

Firstly, we investigated the sharpness of the  $\theta$ -criterion and the accompanying error analysis. For this purpose, the complex-valued field  $G(z_i, z_j) = -m_j/(z_i - z_j)$  was used and we calculated the total force in a system consisting of a million sources (see Figure 5.1 for results and further details of this simulation). Since  $G$  is complex-valued the assumption (2.4) is clearly violated. Cancellation effects therefore implies that relative error estimates may well be impossible to obtain. Nevertheless, since the points are distributed randomly and an irregular mesh is used the effect is negligible in this case (although somewhat more prominent for the uniform distribution, see Figure 5.1).

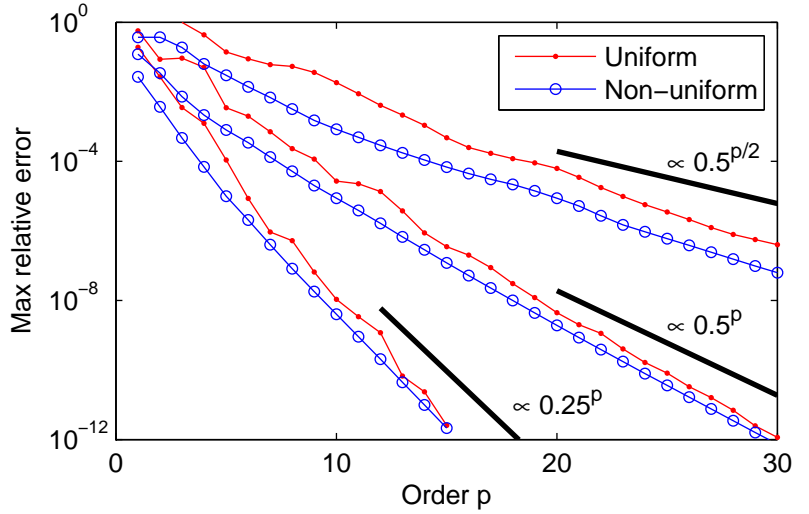


Figure 5.1: Relative error as a function of  $p$  for  $\theta \in \{0.25, 0.5, \sqrt{0.5}\}$  and for two different distributions. A million point sources of total mass 1 are distributed inside the unit circle (i) uniformly, and (ii) non-uniformly with radial density  $\propto 1/r$  (see Figure 1.1 for the resulting mesh in the latter case). For convenience, the error is measured in a random sample ( $M = 1000$ ) of points.

Secondly, we evaluated the efficiency of the proposed adaptivity at an effective relative tolerance of about  $10^{-6}$  (using  $p = 20$  expansion coefficients). Since our code is developed from a highly optimized uniform code outlined in [6], a reasonably fair comparison is possible. Evidently, a square and uniformly subdivided multipole mesh implies a completely regular access pattern so that a highly efficient implementation using direct addressing techniques is possible. It is therefore of some interest to estimate under what circumstances adaptivity is actually beneficial. For this purpose we measured the speedup achieved by the adaptive code when the point sources were sampled from increasingly non-uniform data. The results are displayed in Figure 5.2 and shows that only for very uniform distributions of points is there a small gain ( $\lesssim 15\%$ ) in using the uniform code.

As a third and final experiment we assessed the performance of the adaptive code for different point distributions. Using again a relative tolerance of  $10^{-6}$ , we measured the CPU-time for increasing number of points sampled from three very different distributions. As shown in Figure 5.3, the code is very robust indeed and scales well at least up to some 5 million point sources on a single CPU.

### 5.1. Reproducibility

Our two-dimensional implementation as described in Section 4 is available for download via the author's webpage<sup>2</sup>. The code has been tested on several platforms, is fully documented, and comes with a convenient Matlab mex-interface. Along with the code, automatic Matlab-scripts that repeat the numerical experiments in Section 5 are also distributed. The results presented here were all obtained with a 3.06 GHz Intel Core 2 Duo processor with 4GB of internal memory running under Mac OS X 10.6.6.

<sup>2</sup><http://user.it.uu.se/~stefane/freeware>

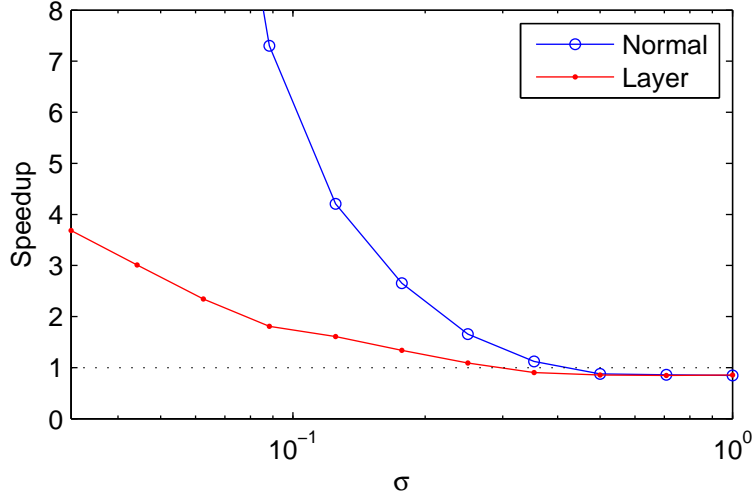


Figure 5.2: The benefits of adaptivity: speedup of the adaptive FMM versus an optimized uniform FMM. The harmonic potential is evaluated at 200,000 random points, chosen either from a normal distribution with variance  $\sigma^2$  or from a combined ‘layer’ distribution where the  $x$ -coordinate is uniformly distributed in  $[0, 1]$ , and the  $y$ -coordinate again is  $N(0, \sigma^2)$ -distributed. For ease of comparison, both distributions are forced by rejection to fit exactly within the unit square. As  $\sigma \rightarrow 0$ , stronger clustering around the origin and the  $x$ -axis occurs, respectively, and the benefits with adaptivity quickly show.

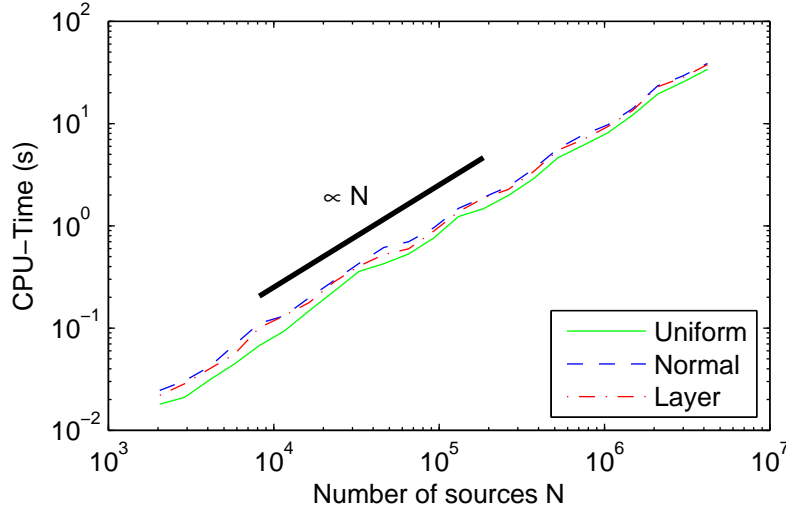


Figure 5.3: Robustness of the adaptive FMM for different distributions of sources. The CPU-time is measured for increasing number of sources and for three different distributions: a uniform distribution in the unit square, and the two distributions used also in Figure 5.2.

## 6. Conclusions

Asymmetric adaptive meshes offer convenient and efficient FMM implementations. If cluster-to-cluster interactions are restricted to boxes for which the  $\theta$ -criterion is true and if the kernel satisfies the assumptions outlined in Section 2, then the relative error can be bounded by a constant  $\times \theta^{p+1}/(1 - \theta)^2$  for  $\theta \in (0, 1)$  and with  $p$  the number of expansion terms. The computational complexity of the resulting 2D-algorithm can be expected to be about  $O(\theta^{-2} \log^{-2} \theta \cdot N \log^2 \text{TOL})$  for some target tolerance TOL. Not only is the actual performance of the algorithm competitive with optimized uniform FMMs even for relatively uniform data, but it is also robust for non-uniform distribution of points. Ongoing work includes porting the code to manycore platforms.

## Acknowledgment

This work was supported by the Swedish Research Council within the FLOW and the UPMARC Linnaeus centers of Excellence.

## References

- [1] S. Aluru, Greengard's  $N$ -body algorithm is not order  $N$ , SIAM J. Sci. Comput. 17 (1996) 773–776. doi:10.1137/S1064827593272031.
- [2] G. Blelloch, G. Narlikar, A practical comparison of  $N$ -body algorithms, in: Parallel Algorithms, volume 30 of *Series in Discrete Mathematics and Theoretical Computer Science*.
- [3] J. Carrier, L. Greengard, V. Rokhlin, A fast adaptive multipole algorithm for particle simulations, SIAM J. Sci. Stat. Comput. 9 (1988) 669–686. doi:10.1137/0909044.
- [4] H. Cheng, L. Greengard, V. Rokhlin, A fast adaptive multipole algorithm in three dimensions, J. Comput. Phys. 155 (1999) 468–498. doi:10.1006/jcph.1999.6355.
- [5] O. Coulaud, P. Fortin, J. Roman, High performance BLAS formulation of the adaptive fast multipole method, Math. Comput. Modelling 51 (2010) 177–188. doi:10.1016/j.mcm.2009.08.039.
- [6] P. Deglaire, S. Engblom, O. Ågren, H. Bernhoff, Analytical solutions for a single blade in vertical axis turbine motion in two-dimensions, Eur. J. Mech. B Fluids 28 (2009) 506–520. doi:10.1016/j.euromechflu.2008.11.004.
- [7] L. Greengard, V. Rokhlin, A fast algorithm for particle simulations, J. Comput. Phys. 73 (1987) 325–348. doi:10.1016/0021-9991(87)90140-9.
- [8] M. Griebel, S. Knapek, G. Zumbusch, Numerical Simulation in Molecular Dynamics, volume 5 of *Texts in Computational Science and Engineering*, Springer Verlag, Berlin, 2007.
- [9] N.A. Gumerov, R. Duraiswami, Fast multipole methods for the Helmholtz equation in three dimensions, Elsevier Series in Electromagnetism, Elsevier, Oxford, 2004.
- [10] N.A. Gumerov, R. Duraiswami, Fast multipole methods on graphics processors, J. Comput. Phys. 227 (2008) 8290–8313. doi:10.1016/j.jcp.2008.05.023.
- [11] H.G. Petersen, E.R. Smith, D. Soelvason, Error estimates for the fast multipole method. II. The three-dimensional case, Proc. Math. Phys. Sci. 448 (1995) 401–418.
- [12] J.K. Salmon, M.S. Warren, Skeletons from the treecode closet, J. Comput. Phys. 111 (1994) 136–155. doi:10.1006/jcph.1994.1050.
- [13] R. Sedgewick, Algorithms in C, Addison-Wesley Series in Computer Science, Addison-Wesley, Reading, MA, 1990.
- [14] B. Shanker, H. Huang, Accelerated Cartesian expansions - a fast method for computing of potentials of the form  $R^{-\nu}$  for all real  $\nu$ , J. Comput. Phys. 226 (2007) 732–753. doi:10.1016/j.jcp.2007.04.033.
- [15] H. Sundar, R.S. Sampath, G. Biros, Bottom-up construction and 2:1 balance refinement of linear octrees in parallel, SIAM J. Sci. Comput. 30 (2008) 2675–2708. doi:10.1137/070681727.
- [16] M. Vikram, A. Baczewski, B. Shanker, S. Aluru, Parallel accelerated Cartesian expansions for particle dynamics simulations, in: Proceedings of the 2009 IEEE International Parallel and Distributed Processing Symposium, pp. 1–11. doi:10.1109/IPDPS.2009.5161038.
- [17] M.S. Warren, J.K. Salmon, A portable parallel particle program, Comput. Phys. Commun. 87 (1995) 266–290. doi:10.1016/0010-4655(94)00177-4.
- [18] J.M. Zhang, M. Tanaka, Adaptive spatial decomposition in fast multipole method, J. Comput. Phys. 226 (2007) 17–28. doi:10.1016/j.jcp.2007.03.032.